2021
ACTIVITY
REPORT

Project-Team
GEOSTAT

Geometry and Statistics in acquisition data

DOMAIN

Applied Mathematics, Computation and Simulation

THEME

Optimization, machine learning and statistical methods

Inria

Contents

Project-Team GEOSTAT

Creation of the Project-Team: 2011 January 01

Keywords

Computer sciences and digital sciences

A3.4.2. – Unsupervised learning

A3.4.7. – Kernel methods

A3.4.8. – Deep learning

A5.3. – Image processing and analysis

A5.3.2. – Sparse modeling and image representation

A5.3.3. – Pattern recognition

A5.3.5. – Computational photography

A5.7. – Audio modeling and processing

A5.7.3. – Speech

A5.7.4. – Analysis

A5.9. – Signal processing

A5.9.2. – Estimation, modeling

A5.9.3. – Reconstruction, enhancement

A5.9.5. – Sparsity-aware processing

Other research topics and application domains

B2. – Health

B2.2. – Physiology and diseases

B2.2.1. – Cardiovascular and respiratory diseases

B2.2.6. – Neurodegenerative diseases

B3. – Environment and planet

B3.3. – Geosciences

B3.3.2. – Water: sea & ocean, lake & river

B3.3.4. – Atmosphere

1   Team members, visitors, external collaborators

Research Scientists

- Hussein Yahia [Team leader, Inria, Researcher, HDR]

- Nicolas Brodu [Inria, Researcher, HDR]

- Khalid Daoudi [Inria, Researcher, HDR]

Post-Doctoral Fellows

- Anis Fradi [Inria, from Nov 2021]

PhD Students

- Biswajit Das [Inria, until Sep 2021]

- Arash Rashidi [Groupe I2S]

Technical Staff

- Zhe Li [Inria, Engineer, from Oct 2021]

- Gabriel Augusto Zebadua Garcia [Inria, Engineer]

Administrative Assistant

- Sabrina Duthil [Inria]

2   Overall objectives

GEOSTAT is a research project which investigates the analysis of some classes of natural complex signals (physiological time series, turbulent universe and earth observation data sets) by determining, in acquired signals, the properties that are predicted by commonly admitted or new physical models best fitting the phenomenon. Consequently, when statistical properties discovered in the signals do not match closely enough those predicted by accepted physical models, we question the validity of existing models or propose, whenever possible, modifications or extensions of existing models. A new direction of research, based on the CONCAUST exploratory action and the newly accepted (in February 2021) associated team COMCAUSA proposed by N. Brodu with USA / UC Davis, Complexity Sciences Center, Physics Department is developed in the team.

An important aspect of the methodological approach is that we don't rely on a predetermined "universal" signal processing model to analyze natural complex signals. Instead, we take into consideration existing approaches in nonlinear signal processing (wavelets, multifractal analysis tools such as log-cumulants or micro-canonical multifractal formalism, time frequency analysis etc.) which are used to determine the micro structures or other micro features inside the acquired signals. Then, statistical analysis of these micro data are determined and compared to expected behaviour from theoretical physical models used to describe the phenomenon from which the data is acquired. From there different possibilities can be contemplated:

- The statistics match behaviour predicted by the model: complexity parameters predicted by the model are extracted from signals to analyze the dynamics of underlying phenomena. Examples: analysis of turbulent data sets in Oceanography and Astronomy.

- The signals displays statistics that cannot be attainable by the common lore of accepted models: how to extend or modify the models according to the behaviour of observed signals ? Example: audio speech signals.

GEOSTAT is a research project in nonlinear signal processing which develops on these considerations: it considers the signals as the realizations of complex extended dynamical systems. The driving approach is to describe the relations between complexity (or information content) and the geometric organization of information in a signal. For instance, for signals which are acquisitions of turbulent fluids, the organization of information may be related to the effective presence of a multiscale hierarchy of coherent structures, of multifractal nature, which is strongly related to intermittency and multiplicative cascade phenomena ; the determination of this geometric organization unlocks key nonlinear parameters and features associated to these signals; it helps understand their dynamical properties and their analysis. We use this approach to derive novel solution methods for super-resolution and data fusion in Universe Sciences acquisitions [12]. Specific advances are obtained in GEOSTAT in using this type of statistical/geometric approach to get validated dynamical information of signals acquired in Universe Sciences, e.g. Oceanography or Astronomy. The research in GEOSTAT encompasses nonlinear signal processing and the study of emergence in complex systems, with a strong emphasis on geometric approaches to complexity. Consequently, research in GEO-STAT is oriented towards the determination, in real signals, of quantities or phenomena, usually unattainable through linear methods, that are known to play an important role both in the evolution of dynamical systems whose acquisitions are the signals under study, and in the compact representations of the signals themselves.

Signals studied in GEOSTAT belong to two broad classes:

- Acquisitions in Astronomy and Earth Observation.

- Physiological time series.

3 Research program

3.1 General methodology

Glossary

Fully Developed Turbulence (FDT) Turbulence at very high Reynolds numbers; systems in FDT are beyond deterministic chaos, and symmetries are restored in a statistical sense only, and multi-scale correlated structures are landmarks. Generalizing to more random uncorrelated multi-scale structured turbulent fields.

Compact Representation Reduced representation of a complex signal (dimensionality reduction) from which the whole signal can be reconstructed. The reduced representation can correspond to points randomly chosen, such as in Compressive Sensing, or to geometric localization related to statistical information content (framework of reconstructible systems).

Sparse representation The representation of a signal as a linear combination of elements taken in a dictionary (frame or Hilbertian basis), with the aim of finding as less as possible non-zero coefficients for a large class of signals.

Universality class In theoretical physics, the observation of the coincidence of the critical exponents (behaviour near a second order phase transition) in different phenomena and systems is called universality. Universality is explained by the theory of the renormalization group, allowing for the determination of the changes followed by structured fluctuations under rescaling, a physical system is the stage of. The notion is applicable with caution and some differences to generalized out-of-equilibrium or disordered systems. Non-universal exponents (without definite classes) exist in some universal slowing dynamical phenomena like the glass transition and kindred. As a consequence, different macroscopic phenomena displaying multiscale structures (and their acquisition in the form of complex signals) may be grouped into different sets of generalized classes.

Every signal conveys, as a measure experiment, information on the physical system whose signal is an acquisition of. As a consequence, it seems natural that signal analysis or compression should make use of physical modelling of phenomena: the goal is to find new methodologies in signal processing that goes beyond the simple problem of interpretation. Physics of disordered systems, and specifically physics of (spin) glasses is putting forward new algorithmic resolution methods in various domains such as optimization, compressive sensing etc. with significant success notably for NP hard problem heuristics. Similarly, physics of turbulence introduces phenomenological approaches involving multifractality. Energy cascades are indeed closely related to geometrical manifolds defined through random processes. At these structures' scales, information in the process is lost by dissipation (close to the lower bound of inertial range). However, all the cascade is encoded in the geometric manifolds, through long or short distance correlations depending on cases. How do these geometrical manifold structures organize in space and time, in other words, how does the scale entropy cascades itself ? To unify these two notions, a description in term of free energy of a generic physical model is sometimes possible, such as an elastic interface model in a random nonlinear energy landscape : This is for instance the correspondence between compressible stochastic Burgers equation and directed polymers in a disordered medium. Thus, trying to unlock the fingerprints of cascade-like structures in acquired natural signals becomes a fundamental problem, from both theoretical and applicative viewpoints.

3.2 Turbulence in insterstellar clouds and Earth observation data

The research described in this section is a collaboration effort of GEOSTAT, CNRS LEGOS (Toulouse), CNRS LAM (Marseille Laboratory for Astrophysics), MERCATOR (Toulouse), IIT Roorkee, Moroccan Royal Center for Teledetection (CRST), Moroccan Center for Science CNRST, Rabat University, University of Heidelberg. Researchers involved:

- GEOSTAT: H. Yahia, N. Brodu, K. Daoudi, A. El Aouni, A. Tamim

- CNRS LAB: S. Bontemps, N. Schneider

- CNRS LEGOS: V. Garçon, I. Hernandez-Carrasco, J. Sudre, B. Dewitte

- CNRST, CRTS, Rabat University: D. Aboutajdine, A. Atillah, K. Minaoui

The analysis and modeling of natural phenomena, specially those observed in geophysical sciences and in astronomy, are influenced by statistical and multiscale phenomenological descriptions of turbulence; indeed these descriptions are able to explain the partition of energy within a certain range of scales. A particularly important aspect of the statistical theory of turbulence lies in the discovery that the support of the energy transfer is spatially highly non uniform, in other terms it is *intermittent* [42]. Because of the absence of localization of the Fourier transform, linear methods are not successful to unlock the multiscale structures and cascading properties of variables which are of primary importance as stated by the physics of the phenomena. This is the reason why new approaches, such as DFA (Detrented Fluctuation Analysis), Time-frequency analysis, variations on curvelets [40] etc. have appeared during the last decades. Recent advances in dimensionality reduction, and notably in Compressive Sensing, go beyond the Nyquist rate in sampling theory using nonlinear reconstruction, but data reduction occur at random places, independently of geometric localization of information content, which can be very useful for acquisition purposes, but of lower impact in signal analysis. We are successfully making use of a microcanonical formulation of the multifractal theory, based on predictability and reconstruction, to study the turbulent nature of interstellar molecular or atomic clouds. Another important result obtained in GEOSTAT is the effective use of multiresolution analysis associated to optimal inference along the scales of a complex system. The multiresolution analysis is performed on dimensionless quantities given by the *singularity exponents* which encode properly the geometrical structures associated to multiscale organization. This is applied successfully in the derivation of high resolution ocean dynamics, or the high resolution mapping of gaseous exchanges between the ocean and the atmosphere; the latter is of primary importance for a quantitative evaluation of global warming. Understanding the dynamics of complex systems is recognized as a new discipline, which makes use of theoretical and methodological foundations coming from nonlinear physics, the study of dynamical systems and many aspects of computer science. One of the challenges is related to the question of *emergence* in complex systems: large-scale effects measurable macroscopically from
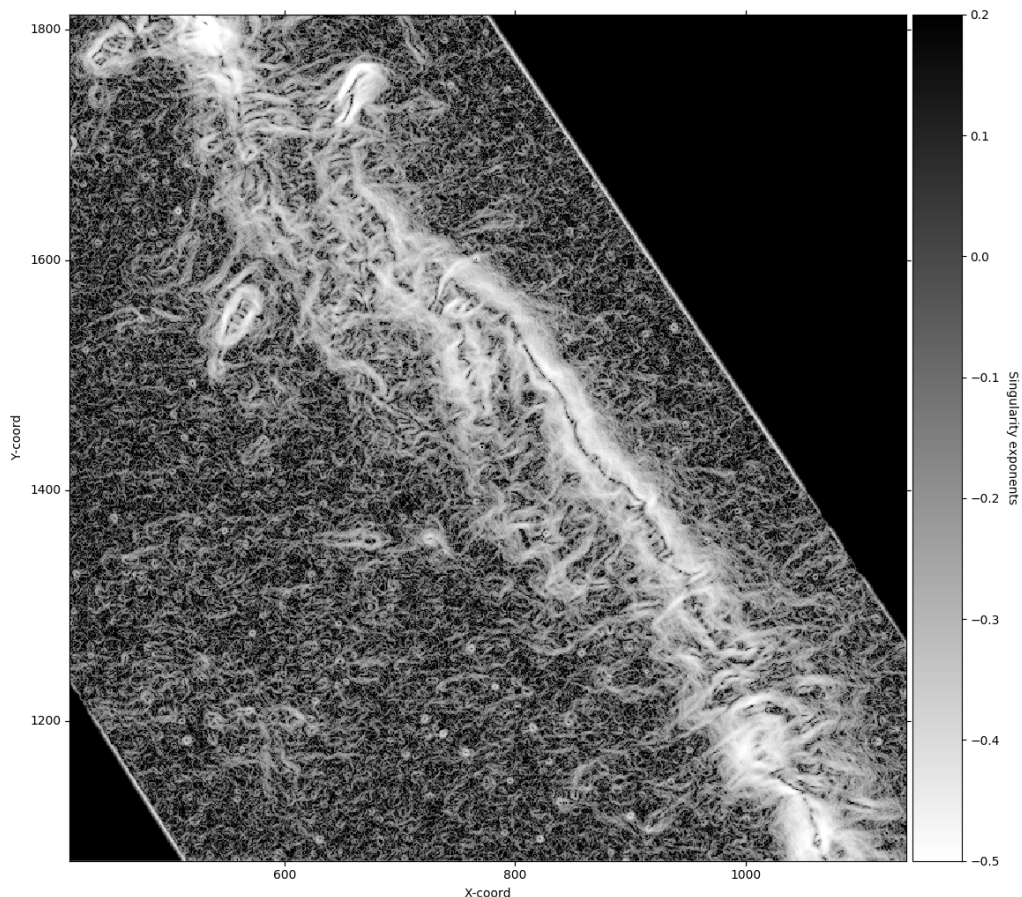
Figure 1: Visualization of the singularity exponents, computed on a edge-aware filtered Musca *Herschel* observation map .

a system made of huge numbers of interactive agents [21, 37]. Some quantities related to nonlinearity, such as Lyapunov exponents, Kolmogorov-Sinai entropy etc. can be computed at least in the phase space [22]. Consequently, knowledge from acquisitions of complex systems (which include *complex signals*) could be obtained from information about the phase space. A result from F. Takens [41] about strange attractors in turbulence has motivated the theoretical determination of nonlinear characteristics associated to complex acquisitions. Emergence phenomena can also be traced inside complex signals themselves, by trying to localize information content geometrically. Fundamentally, in the nonlinear analysis of complex signals there are broadly two approaches: characterization by attractors (embedding and bifurcation) and time-frequency, multiscale/multiresolution approaches. In real situations, the phase space associated to the acquisition of a complex phenomenon is unknown. It is however possible to relate, inside the signal's domain, local predictability to local reconstruction [13] and to deduce relevant information associated to multiscale geophysical signals [14]. A multiscale organization is a fundamental feature of a complex system, it can be for example related to the cascading properties in turbulent systems. We make use of this kind of description when analyzing turbulent signals: intermittency is observed within the inertial range and is related to the fact that, in the case of FDT (fully developed turbulence), symmetry is restored only in a statistical sense, a fact that has consequences on the quality of any nonlinear signal representation by frames or dictionaries.

The example of FDT as a standard "template" for developing general methods that apply to a vast class of complex systems and signals is of fundamental interest because, in FDT, the existence of a multiscale hierarchy $\mathscr{F}_h$ which is of multifractal nature and geometrically localized can be derived from physical considerations. This geometric hierarchy of sets is responsible for the shape of the computed singularity spectra, which in turn is related to the statistical organization of information content in a signal. It explains

scale invariance, a characteristic feature of complex signals. The analogy from statistical physics comes from the fact that singularity exponents are direct generalizations of *critical exponents* which explain the macroscopic properties of a system around critical points, and the quantitative characterization of *universality classes*, which allow the definition of methods and algorithms that apply to general complex signals and systems, and not only turbulent signals: signals which belong to a same universality class share common statistical organization. During the past decades, canonical approaches permitted the development of a well-established analogy taken from thermodynamics in the analysis of complex signals: if $\mathscr{F}$ is the free energy, $\mathscr{T}$ the temperature measured in energy units, $\mathscr{U}$ the internal energy per volume unit $\mathscr{S}$ the entropy and β $1/\mathscr{T}$, then the scaling exponents associated to moments of intensive variables $p \to \tau_p$ corresponds to $\beta\mathscr{F}$, $\mathscr{U}(\beta)$ corresponds to the singularity exponents values, and $\mathscr{S}(\mathscr{U})$ to the singularity spectrum [20]. The research goal is to be able to determine universality classes associated to acquired signals, independently of microscopic properties in the phase space of various complex systems, and beyond the particular case of turbulent data [35].

We show in figure 1 the result of the computation of singularity exponents on an *Herschel* astronomical observation map (the Musca galactic cloud) which has been edge-aware filtered using sparse L^1 filtering to eliminate the cosmic infrared background (or CIB), a type of noise that can modify the singularity spectrum of a signal.

3.3   Causal modeling

The team is working on a new class of models for modeling physical systems, starting from measured data and accounting for their dynamics [27]. The idea is to statistically describe the evolution of a system in terms of causally-equivalent states; states that lead to the same predictions [23]. Transitions between these states can be reconstructed from data, leading to a theoretically-optimal predictive model [39]. In practice, however, no algorithm is currently able to reconstruct these models from data in a reasonable time and without substantial discrete approximations. Recent progress now allows a continuous formulation of predictive causal models. Within this framework, more efficient algorithms may be found. The broadened class of predictive models promises a new perspective on structural complexity in many applications.

3.4   Speech analysis

Phonetic and sub-phonetic analysis: We developed a novel algorithm for automatic detection of Glottal Closure Instants (GCI) from speech signals using the Microcanonical Multiscale Formalism (MMF). This state of the art algorithm is considered as a reference in this field. We made a Matlab code implementing it available to the community (link). Our approach is based on the Microcanonical Multiscale Formalism. We showed that in the case of clean speech, our algorithm performs almost as well as a recent state-of-the-art method. In presence of different types of noises, we showed that our method is considerably more accurate (particularly for very low SNRs). Moreover, our method has lower computational times does not rely on an estimate of pitch period nor any critical choice of parameters. Using the same MMF, we also developed a method for phonetic segmentation of speech signal. We showed that this method outperforms state of the art ones in term of accuracy and efficiency.

Pathological speech analysis and classification: we made a critical analysis of some widely used methodologies in pathological speech classification. We then introduced some novel methods for extracting some common features used in pathological speech analysis and proposed more robust techniques for classification.

Speech analysis of patients with Parkinsonism: with our collaborators from the Czech Republic, we started preliminary studies of some machine learning issues in the field essentially due the small amount of training data.

3.5   Data-based identification of characteristic scales and automated modeling

Data are often acquired at the highest possible resolution, but that scale is not necessarily the best for modeling and understanding the system from which data was measured. The intrinsic properties of natural processes do not depend on the arbitrary scale at which data is acquired; yet, usual analysis techniques operate at the acquisition resolution. When several processes interact at different scales, the identification

of their characteristic scales from empirical data becomes a necessary condition for properly modeling the system. A classical method for identifying characteristic scales is to look at the work done by the physical processes, the energy they dissipate over time. The assumption is that this work matches the most important action of each process on the studied natural system, which is usually a reasonable assumption. In the framework of time-frequency analysis [31], the power of the signal can be easily computed in each frequency band, itself matching a temporal scale.

However, in open and dissipating systems, energy dissipation is a prerequisite and thus not necessarily the most useful metric to investigate. In fact, most natural, physical and industrial systems we deal with fall in this category, while balanced quasi-static assumptions are practical approximation only for scales well below the characteristic scale of the involved processes. Open and dissipative systems are not locally constrained by the inevitable rise in entropy, thus allowing the maintaining through time of mesoscopic ordered structures. And, according to information theory [33], more order and less entropy means that these structures have a higher information content than the rest of the system, which usually gives them a high functional role.

We propose to identify characteristic scales not only with energy dissipation, as usual in signal processing analysis, but most importantly with information content. Information theory can be extended to look at which scales are most informative (e.g. multi-scale entropy [26], ε-entropy [25]). Complexity measures quantify the presence of structures in the signal (e.g. statistical complexity [28], MPR [36] and others [30]). With these notions, it is already possible to discriminate between random fluctuations and hidden order, such as in chaotic systems [27], [36]. The theory of how information and structures can be defined through scales is not complete yet, but the state of art is promising [29]. Current research in the team focuses on how informative scales can be found using collections of random paths, assumed to capture local structures as they reach out [24].

Building on these notions, it should also possible to fully automate the modeling of a natural system. Once characteristic scales are found, causal relationships can be established empirically. They are then clustered together in internal states of a special kind of Markov models called ε-machines [28]. These are known to be the optimal predictors of a system, with the drawback that it is currently quite complicated to build them properly, except for small system [38]. Recent extensions with advanced clustering techniques [23], [32], coupled with the physics of the studied system (e.g. fluid dynamics), have proved that ε-machines are applicable to large systems, such as global wind patterns in the atmosphere [34]. Current research in the team focuses on the use of reproducing kernels, coupled possibly with sparse operators, in order to design better algorithms for ε-machines reconstruction. In order to help with this long-term project, a collaboration is ongoing with J. Crutchfield lab at UC Davis.

4   Application domains

4.1   Sparse signals & optimization

This research topic involves **Geostat** team and is used to set up an InnovationLab with I2S company

Sparsity can be used in many ways and there exist various sparse models in the literature; for instance minimizing the l_0 quasi-norm is known to be an NP-hard problem as one needs to try all the possible combinations of the signal's elements. The l_1 norm, which is the convex relation of the l_0 quasi-norm results in a tractable optimization problem. The l_p pseudo-norms with $0 \ p \ 1$ are particularly interesting as they give closer approximation of l_0 but result in a non-convex minimization problem. Thus, finding a global minimum for this kind of problem is not guaranteed. However, using a non-convex penalty instead of the l_1 norm has been shown to improve significantly various sparsity-based applications. Nonconvexity has a lot of statistical implications in signal and image processing. Indeed, natural images tend to have a heavy-tailed (kurtotic) distribution in certain domains such as wavelets and gradients. Using the l_1 norm comes to consider a Laplacian distribution. More generally, the hyper-Laplacian distribution is related to the l_p pseudo-norm $(0 \ p \ 1)$ where the value of p controls how the distribution is heavy-tailed. As the hyper-Laplacian distribution for $0 \ p \ 1$ represents better the empirical distribution of the transformed images, it makes sense to use the l_p pseudo-norms instead of l_1. Other functions that better reflect heavy-tailed distributions of images have been used as well such as Student-t or Gaussian Scale Mixtures. The internal

properties of natural images have helped researchers to push the sparsity principle further and develop highly efficient algorithms for restoration, representation and coding. Group sparsity is an extension of the sparsity principle where data is clustered into groups and each group is sparsified differently. More specifically, in many cases, it makes sense to follow a certain structure when sparsifying by forcing similar sets of points to be zeros or non-zeros simultaneously. This is typically true for natural images that represent coherent structures. The concept of group sparsity has been first used for simultaneously shrinking groups of wavelet coefficients because of the relations between wavelet basis elements. Lastly, there is a strong relationship between sparsity, nonpredictability and scale invariance.

We have shown that the two powerful concepts of sparsity and scale invariance can be exploited to design fast and efficient imaging algorithms. A general framework has been set up for using non-convex sparsity by applying a first-order approximation. When using a proximal solver to estimate a solution of a sparsity-based optimization problem, sparse terms are always separated in subproblems that take the form of a proximal operator. Estimating the proximal operator associated to a non-convex term is thus the key component to use efficient solvers for non-convex sparse optimization. Using this strategy, only the shrinkage operator changes and thus the solver has the same complexity for both the convex and non-convex cases. While few previous works have also proposed to use non-convex sparsity, their choice of the sparse penalty is rather limited to functions like the l_p pseudo-norm for certain values of $p \geq 0.5$ or the Minimax Concave (MC) penalty because they admit an analytical solution. Using a first-order approximation only requires calculating the (super)gradient of the function, which makes it possible to use a wide range of penalties for sparse regularization. This is important in various applications where we need a flexible shrinkage function such as in edge-aware processing. Apart from non-convexity, using a first-order approximation makes it easier to verify the optimality condition of proximal operator-based solvers via fixed-point interpretation. Another problem that arises in various imaging applications but has attracted less works is the problem of multi-sparsity, when the minimization problem includes various sparse terms that can be non-convex. This is typically the case when looking for a sparse solution in a certain domain while rejecting outliers in the data-fitting term. By using one intermediate variable per sparse term, we show that proximal-based solvers can be efficient. We give a detailed study of the Alternating Direction Method of Multipliers (ADMM) solver for multi-sparsity and study its properties. The following subjects are addressed and receive new solutions:

1. **Edge aware smoothing**: given an input image g, one seeks a smooth image u "close" to g by minimizing:

$$\underset{u}{\operatorname{argmin}} \frac{\lambda}{2} \|u - g\|_2^2 \ \psi(\nabla u)$$

where ψ is a sparcity-inducing non-convex function and λ a positive parameter. Splitting and alternate minimization lead to the sub-problems:

$$
\begin{aligned}
(\text{sp1}) \quad &: \quad v^{(k1)} \leftarrow \underset{v}{\operatorname{argmin}} \ \psi(v) \ \frac{\beta}{2} \|\nabla u^{(k)} - v\|_2^2 \\
(\text{sp2}) \quad &: \quad u^{(k1)} \leftarrow \underset{u}{\operatorname{argmin}} \ \lambda \|u - g\|_2^2 \ \beta \|\nabla u - v^{(k1)}\|_2^2.
\end{aligned}
$$

We solve sub-problem (sp2) through deconvolution and efficient estimation via separable filters and warm-start initialization for fast GPU implementation, and sub-problem (sp1) through non-convex proximal form.

2. **Structure-texture separation**: design of an efficient algorithm using non-convex terms on both the data-fitting and the prior. The resulting problem is solved via a combination of Half-Quadratic (HQ) and Maximization-Minimization (MM) methods. We extract challenging texture layers outperforming existing techniques while maintaining a low computational cost. Using spectral sparsity in the framework of low-rank estimation, we propose to use robust Principal Component Analysis (RPCA) to

perform robust separation on multi-channel images such as glare and artifacts removal of flash/no-flash photographs. As in this case, the matrix to decompose has much less columns than lines, we propose to use a QR decomposition trick instead of a direct singular value decomposition (SVD) which makes the decomposition faster.

3. **Robust integration**: in many applications, we need to reconstruct an image from corrupted gradient fields. The corruption can take the form of outliers only when the vector field is the result of transformed gradient fields (low-level vision), or mixed outliers and noise when the field is estimated from corrupted measurements (surface reconstruction, gradient camera, Magnetic Resonance Imaging (MRI) compressed sensing, etc.). We use non-convexity and multi-sparsity to build efficient integrability enforcement algorithms. We present two algorithms : 1) a local algorithm that uses sparsity in the gradient field as a prior together with a sparse data-fitting term, 2) a non-local algorithm that uses sparsity in the spectral domain of non-local patches as a prior together with a sparse data-fitting term. Both methods make use of a multi-sparse version of the Half-Quadratic solver. The proposed methods were the first in the literature to propose a sparse regularization to improve integration. Results produced with these methods significantly outperform previous works that use no regularization or simple l_1 minimization. Exact or near-exact recovery of surfaces is possible with the proposed methods from highly corrupted gradient fields with outliers.

4. **Learning image denoising**: deep convolutional networks that consist in extracting features by repeated convolutions with high-pass filters and pooling/downsampling operators have shown to give near-human recognition rates. Training the filters of a multi-layer network is costly and requires powerful machines. However, visualizing the first layers of the filters shows that they resemble wavelet filters, leading to sparse representations in each layer. We propose to use the concept of scale invariance of multifractals to extract invariant features on each sparse representation. We build a bi-Lipschitz invariant descriptor based on the distribution of the singularities of the sparsified images in each layer. Combining the descriptors of each layer in one feature vector leads to a compact representation of a texture image that is invariant to various transformations. Using this descriptor that is efficient to calculate with learning techniques such as classifiers combination and artificially adding training data, we build a powerful texture recognition system that outperforms previous works on 3 challenging datasets. In fact, this system leads to quite close recognition rates compared to latest advanced deep nets while not requiring any filters training.

5   Social and environmental responsibility

5.1   Participation in the Covid-19 Inria mission

GeoStat is participating in the Covid-19 Inria mission: : Vocal biomarkers of respiratory diseases.

6   Highlights of the year

6.1   Awards

- The article *Description of turbulent dynamics in the interstellar medium: multifractal/microcanonical analysis I. Application to Herschel observations of the Musca filament* has been selected as *Highlights* by Astronomy and Astrophysics: link to the *Highlights* section of Astronomy and Astrophysics. The corresponding research topic has also been the subject of an INRIA article on inria.fr: link to to the INRIA article.

- Research-business partnership in image processing: report between I2S and Geostat. Article published on inria.fr: link to to the INRIA article.

6.2   HDRs

- Nicolas Brodu defended his HDR entitled *Systèmes complexes : inférence, dynamique et applications* on June 9, 2021.

- K. Daoudi defended his HDR entitled *Nouveaux paradigmes en traitement de la parole et de ses troubles* on March 17, 2021.

7   New software and platforms

7.1   New software

7.1.1   Fluex

Keywords: Signal, Signal processing

Scientific Description: Fluex is a package consisting of the Microcanonical Multiscale Formalism for 1D, 2D 3D and 3D+t general signals.

Functional Description: Fluex is a C++ library developed under Gforge. Fluex is a library in nonlinear signal processing. Fluex is able to analyze turbulent and natural complex signals, Fluex is able to determine low level features in these signals that cannot be determined using standard linear techniques.

Contact: Hussein Yahia

Participants: Hussein Yahia, Rémi Paties

7.1.2   FluidExponents

Keywords: Signal processing, Wavelets, Fractal, Spectral method, Complexity

Functional Description: FluidExponents is a signal processing software dedicated to the analysis of complex signals displaying multiscale properties. It analyzes complex natural signals by use of nonlinear methods. It implements the multifractal formalism and allows various kinds of signal decomposition and reconstruction. One key aspect of the software lies in its ability to evaluate key concepts such as the degree of impredictability around a point in a signal, and provides different kinds of applications. The software can be used for times series or multidimensional signals.

Contact: Hussein Yahia

Participants: Antonio Turiel, Hussein Yahia

7.1.3   ProximalDenoising

Name: ProximalDenoising

Keywords: 2D, Image filter, Filtering, Minimizing overall energy, Noise, Signal processing, Image reconstruction, Image processing

Scientific Description: Image filtering is contemplated in the form of a sparse minimization problem in a non-convex setting. Given an input image I, one seeks to compute a denoised output image u such that u is close to I in the L2 norm. To do so, a minimization term is added which favors sparse gradients for output image u. Imposing sparse gradients lead to a non-convex minimization term: for instance a pseudo-norm Lp with $0 < p < 1$ or a Cauchy or Welsh function. Half-quadratic algorithm is used by adding a new variable in the minimization functionnal which leads to two sub-problems, the first sub-problem is non-convex and solved by use of proximal operators. The second sub-problem can be written in variational form, and is best solved in Fourier space: it takes the form of a deconvolution operator whose kernel can be approximated by a finite sum of separable filters. This solution method produces excellent computation times even on big images.

Functional Description: Use of proximal and non quadratic minimization. GPU implementation. If f is an input image, one seeks an output g such that the following functional is minimized:

l/2*(norme2(f-g) + psi(grad(g))) with : l positive constant, norme2 = L2 norm, psi is a Cauchy function used for parcimony.

This functional is also applied for debayerization.

Release Contributions: This software implements H. Badri PhD thesis results.

URL: https://gitlab.inria.fr/marmarti/i2s_geostat_C

Authors: Marie Martin, Chiheb Sakka, Hussein Yahia, Nicolas Brodu, Gabriel Zebadua Garcia, Khalid Daoudi

Contact: Hussein Yahia

Partner: Innovative Imaging Solutions I2S

7.1.4 Manzana

Name: Manzana

Keywords: 2D, Image processing, Filtering

Scientific Description: Software library developed in the framework of I2S-GEOSTAT innovationlab and made of high-level image processing functionalities based on sparsity and non-convex optimization.

Functional Description: Library of software in image processing: filtering, hdr, inpainting etc.

Contact: Hussein Yahia

Partner: Innovative Imaging Solutions I2S

7.2 New platforms

Participants: .

8 New results

8.1 Turbulent dynamics in the interstellar medium

Participants: H. Yahia, N. Schneider, S. Bontemps, L. Bonne, et al.

Observations of the interstellar medium (ISM) show a complex density and velocity structure, which is in part attributed to turbulence. Consequently, the multifractal formalism should be applied to observation maps of the ISM in order to characterize its turbulent and multiplicative cascade properties. However, the multifractal formalism, even in its more advanced and recent canonical versions, requires a large number of realizations of the system, which usually cannot be obtained in astronomy. We present a self-contained introduction to the multifractal formalism in a "microcanonical" version, which allows us, for the first time, to compute precise turbulence characteristic parameters from a single observational map without the need for averages in a grand ensemble of statistical observables (e.g., a temporal sequence of images). We compute the singularity exponents and the singularity spectrum for both observations and magnetohydrodynamic simulations, which include key parameters to describe turbulence in the ISM. For the observations we focus on the 250 μm Herschel map of the Musca filament. Scaling properties are investigated using spatial 2D structure functions, and we apply a two-point log-correlation magnitude analysis over various lines of the spatial observation, which is known to be directly related to the existence of a multiplicative cascade under precise conditions. It reveals a clear signature of a multiplicative cascade in Musca with an inertial range from 0.05 to 0.65 pc. We show that the proposed microcanonical approach provides singularity spectra that are truly scale invariant, as required to validate any method used to analyze multifractality. The obtained

singularity spectrum of Musca, which is sufficiently precise for the first time, is clearly not as symmetric as usually observed in log-normal behavior. We claim that the singularity spectrum of the ISM toward Musca features a more log-Poisson shape. Since log-Poisson behavior is claimed to exist when dissipation is stronger for rare events in turbulent flows, in contrast to more homogeneous (in volume and time) dissipation events, we suggest that this deviation from log-normality could trace enhanced dissipation in rare events at small scales, which may explain, or is at least consistent with, the dominant filamentary structure in Musca. Moreover, we find that subregions in Musca tend to show different multifractal properties: While a few regions can be described by a log-normal model, other regions have singularity spectra better fitted by a log-Poisson model. This strongly suggests that different types of dynamics exist inside the Musca cloud. We note that this deviation from log-normality and these differences between subregions appear only after reducing noise features, using a sparse edge-aware algorithm, which have the tendency to "log-normalize" an observational map. Implications for the star formation process are discussed. Our study establishes fundamental tools that will be applied to other galactic clouds and simulations in forthcoming studies.

Publication: Astronomy & Astrophysics, HAL. This publication has been selected in the "Highlights" section of Astronomy & Astrophysics.

8.2 CONCAUST Exploratory Action

Participants: N. Brodu, J. P. Crutchfield, L. Bourel, P. Rau, A. Rupe, Y. Li.

Concaust Exploratory Action Web page.

The article detailing the exploratory action core method "Discovering Causal Structure with Reproducing-Kernel Hilbert Space ε-Machines", has been accepted for publication in the Chaos journal. The remaining budget was reallocated into 2 years of post-doctorate positions: - 1 year for working on the core methodology - 1 year for researching the applicability of the method on real data, in particular the study of El Niño phenomena. Due to covid constraints, the postdoctorate candidate (Alexandra Jurgens) that was selected for the core methodology could not come to France in fall 2021. She may come again in fall 2022, if the situation allows, so the corresponding post-doc is reserved for her. The second post-doctorate offer on El Niño is currently opened for applications.

8.3 COMCAUSA Associated team

Participants: N. Brodu, J. P. Crutchfield, et al. (see Website).

Web page The associate team Comcausa was created as part of the Inria@SiliconValley international lab, between Inria Geostat and the Complexity Sciences Center at University of California, Davis. This team is managed by Nicolas Brodu (Inria) and Jim Crutchfield (UC Davis) and the full list of collaborators is given on the web site. We organized a series of 10 online seminars "Inference for Dynamical Systems A seminar series" in which we invited team members and external researchers to present their results. This online seminars series was a federative moment during the covid lockdowns and fostered new collaborations. Additional funding was obtained (co-PIs Nicolas Brodu, Jim Crutchfield, Sarah Marzen) from the Templeton Foundation in the form of 2×1 years post-doctorate, to work on bioacoustic signatures in whale communication signals. We recruited Alexandra Jurgens for one year, renewable, on an exploratory topic of research: seeking new methods for inferring how much information is being transferred at every scale in a signal. Nicolas Brodu is actively co-supervising her on this program, which she may pursue at Inria in fall 2022 on the Concaust exploratory action post-doc budget. More preliminary results from this associate team were obtained on CO and water flux in$_2$ ecosystems (collaboration between Nicolas Brodu, Yao Liu and Adam Rupe). Nicolas Brodu presented these at the yearly meeting of the ICOS network on monitoring stations, run mostly by INRAE. This in turn lead to the writing of a proposal for the joint Inria-INRAE « Agroécologie et numérique » PEPR, which passed the pre-selection phase in December 2021: this project is being proposed, jointly with a partner at INRAE, as one of the 10 flagship projects retained for the round 1 of this PEPR. The

final decision for whether this PEPR will be funded or not will be made in 2022. Similarly, preliminary results from the El Niño data (collaboration between Nicolas Brodu and Luc Bourrel), have lead to the submission of an ANR proposal. This ANR funding would allow us to extend the work of the post-doctorate researcher which we will co-supervise on the Concaust budget. The Associated Team budget of 2021 could only be partially used as travels were restricted for most of the year. An extensive lab tour was still possible (Nov.-Dec. 2021), where Nicolas Brodu has met with most US associate team members. This tour was scientifically fruitful and we are currently preparing articles detailing our new results

8.4 Vocal biomarkers of Parkinsonian disorders

Participants: K. Daoudi, B. Das, S. Milhé de Saint Victor, A. Foubert-Samier, A. Pavy-Le Traon, O. Rascol, W. G. Meissner, V. Woisard.

In the early stage of disease, the symptoms of Parkinson's disease (PD) are similar to atypical Parkinsonian disorders (APD) such as Progressive Supranuclear Palsy (PSP) and Multiple System Atrophy (MSA). The early differential diagnosis between PD and APD and between APD groups is thus a very challenging task. It turns out that speech disorder is an early and common symptom to PD and APD. The goal of our research is to develop a digital biomarker based on speech analysis in order to assist the neurologists in their diagnosis. We identified distinctive speech features for discrimination between PD and MSA-P, the variant of MSA where Parkinsonism dominates. These features were inferred from French speaking patients and consist in the detection of distortions in the production of particular voiced consonants (plosives and fricatives). We also continued of work on differential diagnosis between PSP and MSA Czech speaking patients. We designed two composite speech indices based on two categorizations of speech features, production subsystems and dysarthria type. These new indices led to high accuracy discrimination not only between PSP and MSA but also between MSA/PSP and PD. These results made us confident to step on to the second clinical phase of the Voice4PD-MSA project, the validation phase, thought we didn't reach the expected number of patient inclusions in the first phase (because of the pandemic). The clinical protocol of this second phase will be submitted to CPP in January 2022.

Publications: PhD thesis (B. Das HAL), Proc. Interspeech' 2021 (K. Daoudi *et al.* HAL).

8.5 Vocal biomarkers of respiratory diseases

Participants: K. Daoudi, T. Similowski.

GeoStat continues its contribution to the Covid-19 mission of Inria via the VocaPnée project in partnership with AP-HP and co-directed by K. Daoudi and Thomas Similowski, responsible for the pulmonology and resuscitation service at La Pitié-Salpêtrière hospital and UMR-S 1158. The objective of the VocaPnée project is to bring together skills available at Inria to develop and validate a vocal biomarker for the remote monitoring of patients at home suffering from acute respiratory diseases (such as Covid-19) or chronic one (such as asthma). This biomarker will then be integrated into a telemedicine platform, ORTIF or COVIDOM for Covid, to assist the doctors in assessing the patient's respiratory status. VocaPnée is divided into 2 longitudinal clinical studies, a hospital study, as a proof of concept, followed by another in tele-medicine. The former received the clearance of CPP and we are preparing the file for the CNIL clearance. In this context, the ADT project VocaPy started in November for a 2 years duration. The goal of VocaPy is to develop a Python library dedicated to pathological speech analysis and vocal biomarkers conception. VocaPy works in synergy with the ADT VocaPnée-Infra dedicated to the development of an infrastructure to interact with protected health data.

8.6 Bayesian approach in a learning-based hyperspectral image denoising framework

Participants: H. Aetesam, S. Kumar Maji, H. Yahia.

Hyperspectral images are corrupted by a combination of Gaussian-impulse noise. On one hand, the traditional approach of handling the denoising problem using maximum a posteriori criterion is often restricted by the time-consuming iterative optimization process and design of hand-crafted priors to obtain an optimal result. On the other hand, the discriminative learning-based approaches offer fast inference speed over a trained model; but are highly sensitive to the noise level used for training. A discriminative model trained with a loss function which does not accord with the Bayesian degradation process often leads to sub-optimal results. In this paper, we design the training paradigm emphasizing the role of loss functions; similar to as observed in model-based optimization methods. As a result; loss functions derived in Bayesian setting and employed in neural network training boosts the denoising performance. Extensive analysis and experimental results on synthetically corrupted and real hyperspectral dataset suggest the potential applicability of the proposed technique under a wide range of homogeneous and heterogeneous noisy settings.

Publication: IEEE Access , HAL

8.7   InnovationLab with I2S

Participants: A. Zebadua, A. Rashidi, H. Yahia, A. Cherif *(I2S)*, J. L. Val-
lancogne *(I2S)*, A. Cailly *(I2S)*.

The InnovationLab with I2S is extended one year starting 1st February 2021.

8.8   InnovationLab with I2S:Acceleration of optimization-based methods using convolutional neural networks (Emulation)

Participants: A. Zebadua, H. Yahia, A. Cherif *(I2S)*, J. L. Vallancogne *(I2S)*,
A. Cailly *(I2S)*.

During the Inria-i2s Innovationlab partnership (2017-2020), several iterative optimization-based image processing algorithms were developed. Such methods are able to provide significant image quality improvements with respect to the previous i2s image processing methods. However, such algorithms seek the minimum of a mathematical function iteratively. In some cases, their convergence can be significantly slow and demand an execution time prohibitive for production goals.

This motivated a deep-learning approach in order to reduce computational time. We called this approach Emulation. A convolutional neural network (CNN) is designed and trained it to reproduce the results obtained by the optimization approach. This makes it possible to take advantage of the quality level of an expensive iterative optimization algorithm at a lower computational cost and shorter execution time. Iterative and deep learning methods make both use of numerical optimization to solve a mathematical problem. In the case of deep learning, however, the optimization takes places during training, which is performed offline. Once trained, a CNN can be used online to provide a fast execution time to process an image (which is commonly referred to as inference). This approach was shown to be effective and allowed comparable image quality levels along with a reduction in execution time 5 times shorter.

8.9   InnovationLab with I2S: Fast image deconvolution for enhancement of the resolution in the video rate terahertz imaging

Participants: A. Rashidi, A. Cailly *(I2S)*, A. Cherif *(I2S)*, H. Yahia.
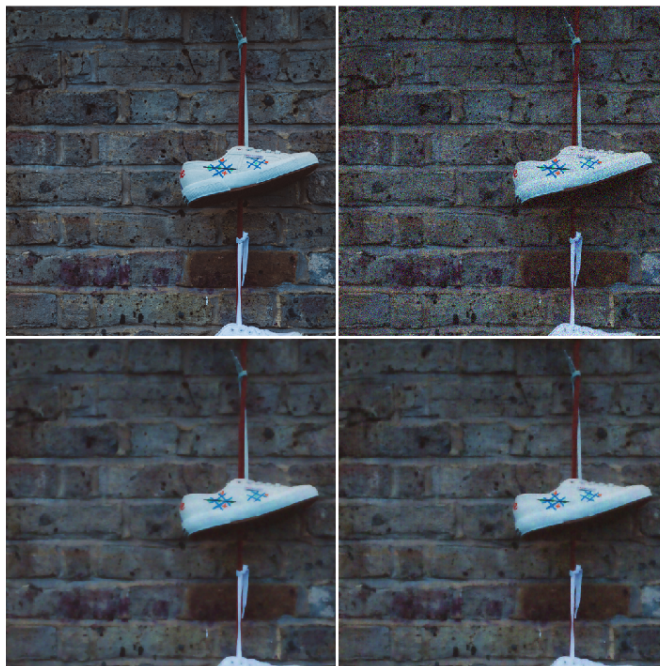
Figure 2: Comparable image-reconstruction quality can be obtained by using Emulation approach. Top right: Original. Top left: corrupted (PSNR :23.539, SSIM: 0.569). Down Left Optimization (PSNR :28.61, SSIM: 0.776). Down Right: Emulation (PSNR :28.8, SSIM: 0.7823). Computation time is 5 times faster with a CNN.

A. Rashidi has pursued the implementation of deconvolution and denoising algorithms within the algorithms of I2S. A fast image deconvolution algorithm is used to demonstrate the resolution enhancement of video rate camera acquired Terahertz images. Our algorithm is based on variable splitting technique with the use of a family of sparsity inducing regularizers for the first time in an image deconvolution application, it is also suitable for practical applications in industry with computationally constrained conditions. The results of the proposed process provide substantial enhancement on the quality and resolution of THz images.

A first work concerned the classical cameras of I2S which is called "Eagle".

Industrial Quality assessment of the images and results showed effective and encouraging results. The images showed better image quality contrast.

Following on, the super-resolution algorithm was integrated into the new cameras of I2S. The new camera is called "Xtra". we are currently in the stage of verifying the effectiveness of the results and their quality with standard measurement methods.

Publication: 46th International Conference on Infrared, Millimeter, and Terahertz Waves, HAL

9 Bilateral contracts and grants with industry

9.1 Bilateral contracts with industry

InnovationLab with I2S company, starting scheduled after 1st 2019 COPIL in January 2019. This InnovationLab is extended one year starting February 2021.

Figure 3: Top: Original image. Center: Denoised using i2s method. Down: Emulation. Higher image quality can be obtained by using an emulation approach which provide an image quality compared to that of an iterative algorithm.

10　Partnerships and cooperations

10.1　International initiatives

10.1.1　Associate Teams in the framework of an Inria International Lab or in the framework of an Inria International Program

COMCAUSA

Title: Computation of causal structures

Duration: 2021 ->

Coordinators: J. P. Crutchfield (chaos@cse.ucdavis.edu), N. Brodu

Partners:

- University of California Davis

Inria contact: N. Brodu

Summary: The associate team Comcausa was created as part of the Inria@SiliconValley international lab, between Inria

10.2　European initiatives

10.2.1　FP7 & H2020 projects

GENESIS Project (Geostat, Laboratoire d'Astrophysique de Bordeaux, Physics Inst. (Köln University) 5-year contract, 2017-2022 (initially 3 years, extented).

Web site. GENeration and Evolution of Structures in the ISm. The objective of this research project is to better understand the structure and evolution of molecular clouds in the interstellar medium (ISM) and to link cloud structure with star-formation. For that, far-infrared observations of dust (Herschel) and cooling lines (SOFIA) are combined with ground-based submillimetre observations of molecular lines. Dedicated analysis tools will be used and developed to analyse the maps and compared to simulations in order to disentangle the underlying physical processes such as gravity, turbulence, magnetic fields, and radiation.

10.3   National initiatives

- CONCAUST Exploratory Action The exploratory action « TRACME » was renamed « CONCAUST » and is going on with good progress. Collaboration with J. P. Crutchfield and its laboratory has lead to a first article, "Discovering Causal Structure with Reproducing-Kernel Hilbert Space ε-Machines", available available here . That article poses the main theoretical fundations for building a new class of models, able to reconstruct a measured process « causal states » from data.

- CovidVoice project: Inria Coind-19 mission. The CovidVoice project evolved into the VocaPnée project in partnership with AP-HP and co-directed by K. Daoudi and Thomas Similowski, responsible for the pulmonology and resuscitation service at La Pitié-Salpêtrière hospital and UMR-S 1158. The objective of the VocaPnée project is to bring together all the skills available at Inria to develop and validate a vocal biomarker for the remote monitoring of patients at home suffering from an acute respiratory disease (such as Covid) or chronic (such as asthma) . This biomarker will then be integrated into a telemedicine platform, ORTIF or COVIDOM for Covid, to assist the doctors in assessing the patient's respiratory status. VocaPnée is divided into 2 longitudinal pilot clinical studies, a hospital study and another in tele-medicine.

  In this context, a voice data collection platform was developed by Inria's SED. This platform is used to collect data from healthy controls. It will then be migrated to the AP-HP servers to collect patient data.

- ANR project *Voice4PD-MSA*, led by K. Daoudi, which targets the differential diagnosis between Parkinson's disease and Multiple System Atrophy. The total amount of the grant is 468555 euros, from which GeoStat has 203078 euros. Initial duration of the project was 42 months, and has been extended until January 2023. Partners: CHU Bordeaux (Bordeaux), CHU Toulouse, IRIT, IMT (Toulouse).

- GEOSTAT is a member of ISIS (Information, Image & Vision), AMF (Multifractal Analysis) GDRs.

- GEOSTAT is participating in the CNRS IMECO project *Intermittence multi-échelles de champs océaniques : analyse comparative d'images satellitaires et de sorties de modèles numériques*. CNRS call AO INSU 2018. PI: F. Schmitt, DR CNRS, UMR LOG 8187. Duration: 2 years.

10.4   Regional initiatives

The ADT (IA Plan) project VocaPy, led by K. Daoudi, started in November 2021 for a 2 years duration. The goal of VocaPy is to develop a Python library dedicated to pathological speech analysis and vocal biomarkers conception. Ms Zhe Li was recruited as an engineer on this project.

11   Dissemination

11.1   Promoting scientific activities

11.1.1   Journal

Member of the editorial boards   H. Yahia is a member of the editorial board of "Frontiers in Fractal Physiology" journal.

11.1.2   Invited talks

- N. Brodu: The ICOS days, Université de Reims, France, 12-14 October 2021

11.1.3 Scientific expertise

- K. Daoudi was invited by ARS Rhône-Alpes (on December 3th 2021) to give a talk on ethical risks of vocal biomarkers: link

- N. Brodu. Santa Fe - Nov 29/Dec 7: Meeting with Jim Crutchfield (Santa Fe Institute, UC Davis) and Adam Rupe (Los Alamos National Lab).

- N. Brodu. Los Alamos - Dec 7/8: Meeting with Balasubramanya Nadiga (Los Alamos National Lab).

- N. Brodu. Santa Barbara - Dec 8/9: Meeting with Nina Miolane (UC Santa Barbara).

- N. Brodu. Pasadena - 9/12: Meeting with Stuart Bartlett (Caltech).

- N. Brodu. Claremont - 12/15: Meeting with Sarah Marzen (Claremont College), Alexandra Jurgens (UC Davis), Alec Boyd (Caltech).

- N. Brodu: European Conference on Complex Systems, Lyon, 25-29 October 2021

11.1.4 Juries

H. Yahia was a member of the PhD thesis jury of H. MAHAMAT. Thesis defended on March 30, 2021, 10.30 am, Université de Bourgogne, UFR Sciences et Techniques - BP 47870 21078 Dijon cedex.

11.2 Popularization

11.2.1 Articles and contents

- INRIA article: H. Yahia: Uncovering the turbulent origins of stars through mathematics.

- INRIA article: H. Yahia: Partneship research-industry in image processing: I2S-Geostat Innovationlab report.

12 Scientific production

12.1 Major publications

[1] G. Attuel, E. Gerasimova-Chechkina, F. Argoul, H. Yahia and A. Arnéodo. 'Multifractal Desynchronization of the Cardiac Excitable Cell Network During Atrial Fibrillation. II. Modeling'. In: *Frontiers in Physiology* 10 (Apr. 2019), 480 (1–18). DOI: 10.3389/fphys.2019.00480. URL: https://hal.inria.fr/hal-02108521.

[2] G. Attuel, E. G. Gerasimova-Chechkina, F. Argoul, H. Yahia and A. Arnéodo. 'Multifractal desynchronization of the cardiac excitable cell network during atrial fibrillation. I. Multifractal analysis of clinical data'. In: *Frontiers in Physiology* 8 (Mar. 2018), pp. 1–30. DOI: 10.3389/fphys.2017.01139. URL: https://hal.inria.fr/hal-01673364.

[3] H. Badri and H. Yahia. 'A Non-Local Low-Rank Approach to Enforce Integrability'. In: *IEEE Transactions on Image Processing* (June 2016), p. 10. URL: https://hal.inria.fr/hal-01317151.

[4] H. Badri, H. Yahia and K. Daoudi. 'Fast and Accurate Texture Recognition with Multilayer Convolution and Multifractal Analysis'. In: *European Conference on Computer Vision*. ECCV 2014. Zürich, Switzerland, Sept. 2014. URL: https://hal.inria.fr/hal-01064793.

[5] I. Hernandez-Carrasco, V. Garçon, J. Sudre, C. Garbe and H. Yahia. 'Increasing the Resolution of Ocean pCO$_2$ Maps in the South Eastern Atlantic Ocean Merging Multifractal Satellite-Derived Ocean Variables'. In: *IEEE Transactions on Geoscience and Remote Sensing* (June 2018), pp. 1–15. DOI: 10.1109/TGRS.2018.2840526. URL: https://hal.inria.fr/hal-01825810.

[6]   I. Hernández-Carrasco, J. Sudre, V. Garçon, H. Yahia, C. Garbe, A. Paulmier, B. Dewitte, S. Illig, I. Dadou, M. González-Dávila and J. Santana Casiano. 'Reconstruction of super-resolution ocean pCO 2 and air-sea fluxes of CO 2 from satellite imagery in the Southeastern Atlantic'. In: *Biogeosciences* (Sept. 2015), p. 20. URL: https://hal.inria.fr/hal-01193339.

[7]   V. Khanagha, K. Daoudi and H. Yahia. 'Detection of Glottal Closure Instants based on the Micro-canonical Multiscale Formalism'. In: *IEEE Transactions on Audio, Speech and Language Processing* (Dec. 2014). URL: https://hal.inria.fr/hal-01059345.

[8]   V. Khanagha, K. Daoudi and H. Yahia. 'Efficient and robust detection of Glottal Closure Instants using Most Singular Manifold of speech signals'. In: *IEEE Transactions on Acoustics Speech and Signal Processing* (2014). forthcoming. URL: https://hal.inria.fr/hal-00802014.

[9]   S. Maji, H. Yahia and T. Fusco. 'A Multifractal-based Wavefront Phase Estimation Technique for Ground-based Astronomical Observations'. In: *IEEE Transactions on Geoscience and Remote Sensing* (Nov. 2015), p. 11. DOI: 10.1109/TGRS.2015.2487546. URL: https://hal.inria.fr/hal-01254482.

[10] S. K. Maji and H. Yahia. 'Edges, Transitions and Criticality'. In: *Pattern Recognition* (Jan. 2014). URL: http://hal.inria.fr/hal-00924137.

[11] O. Pont, A. Turiel and H. Yahia. 'Singularity analysis in digital signals through the evaluation of their Unpredictable Point Manifold'. In: *International Journal of Computer Mathematics* (2012). URL: http://hal.inria.fr/hal-00688715.

[12] J. Sudre, H. Yahia, O. Pont and V. Garçon. 'Ocean Turbulent Dynamics at Superresolution From Optimal Multiresolution Analysis and Multiplicative Cascade'. In: *IEEE Transactions on Geoscience and Remote Sensing* 53.11 (June 2015), p. 12. DOI: 10.1109/TGRS.2015.2436431. URL: https://hal.inria.fr/hal-01166170.

[13] A. Turiel, H. Yahia and C. Perez-Vicente. 'Microcanonical multifractal formalism: a geometrical approach to multifractal systems. Part I: singularity analysis'. In: *Journal of Physics A: Math. Theor* 41 (2008). URL: http://dx.doi.org/10.1088/1751-8113/41/1/015501.

[14] H. Yahia, J. Sudre, C. Pottier and V. Garçon. 'Motion analysis in oceanographic satellite images using multiscale methods and the energy cascade'. In: *Pattern Recognition* 43 (10 2010), pp. 3591–3604. URL: http://dx.doi.org/10.1016/j.patcog.2010.04.011.

12.2   Publications of the year

International journals

[15] H. Aetesam, S. K. Maji and H. Yahia. 'Bayesian Approach in a Learning-Based Hyperspectral Image Denoising Framework'. In: *IEEE Access* (22nd Dec. 2021), pp. 1–1. DOI: 10.1109/ACCESS.2021.3137656. URL: https://hal.inria.fr/hal-03502378.

[16] H. Yahia, N. Schneider, S. Bontemps, L. Bonne, G. Attuel, S. Dib, V. Ossenkopf-Okada, A. Turiel, A. Zebadua, D. Elia, S. Kumar Maji, F. G. Schmitt and J.-F. Robitaille. 'Description of turbulent dynamics in the interstellar medium: Multifractal-microcanonical analysis: I. Application to Herschel observations of the Musca filament'. In: *Astronomy and Astrophysics - A&A* (17th Mar. 2021). DOI: 10.1051/0004-6361/202039874. URL: https://hal.inria.fr/hal-03193056.

Conferences without proceedings

[17] K. Daoudi, B. Das, S. Milhé De Saint Victor, A. Foubert-Samier, A. Pavy-Le Traon, O. Rascol, W. G. Meissner and V. Woisard. 'Distortion of voiced obstruents for differential diagnosis between parkinson's disease and multiple system atrophy'. In: INTERSPEECH 2021. Brno, Czech Republic, 30th Aug. 2021. URL: https://hal.archives-ouvertes.fr/hal-03313375.

[18] A. Rashidi, A. Minasyan, A. Cailly, M. Hamdi, O. Redon, L. Dussopt and H. Yahia. 'Fast image deconvolution for enhancement of the resolution in the video rate terahertz imaging'. In: *IRMMW-THz46th International Conference on Infrared, Millimeter, and Terahertz Waves*. IRMMW-THz. Beijing, China, 3rd Sept. 2021. URL: https://hal.inria.fr/hal-03343739.

Doctoral dissertations and habilitation theses

[19] B. Das. 'Automated objective speech markers for differential diagnosis between parkinson's disease and atypical parkinsonian disorders'. Université de Bordeaux, 30th Sept. 2021. URL: https://tel.archives-ouvertes.fr/tel-03436409.

12.3 Cited publications

[20] A. Arnéodo, F. Argoul, E. Bacry, J. Elezgaray and J. F. Muzy. *Ondelettes, multifractales et turbulence*. Paris, France: Diderot Editeur, 1995.

[21] N. Boccara. *Modeling Complex Systems*. New-York Dordrecht Heidelberg London: Springer, 2010.

[22] G. Boffetta, M. Cencini, M. Falcioni and A. Vulpiani. 'Predictability: a way to characterize complexity'. In: *Physics Report* 356 (2002). arXiv:nlin/0101029v1, pp. 367–474. URL: http://dx.doi.org/10.1016/S0370-1573(01)00025-4.

[23] N. Brodu. 'Reconstruction of epsilon-machines in predictive frameworks and decisional states'. In: *Advances in Complex Systems* 14.05 (2011), pp. 761–794. DOI: 10.1142/S0219525911003347. eprint: https://doi.org/10.1142/S0219525911003347. URL: https://doi.org/10.1142/S0219525911003347.

[24] N. Brodu and H. Yahia. 'Stochastic Texture Difference for Scale-Dependent Data Analysis'. In: *arXiv preprint arXiv:1503.03278* (2015).

[25] P. Castiglione, M. Falcioni, A. Lesne and A. Vulpiani. *Chaos and coarse graining in statistical mechanics*. Cambridge University Press Cambridge, 2008.

[26] M. Costa, A. L. Goldberger and C.-K. Peng. 'Multiscale entropy analysis of complex physiologic time series'. In: *Physical review letters* 89.6 (2002), p. 068102.

[27] J. P. Crutchfield. 'Between order and chaos'. In: *Nature Physics* 8.1 (2012), pp. 17–24.

[28] J. P. Crutchfield and K. Young. 'Inferring statistical complexity'. In: *Physical Review Letters* 63.2 (1989), p. 105.

[29] G. Falasco, G. Saggiorato and A. Vulpiani. 'About the role of chaos and coarse graining in statistical mechanics'. In: *Physica A: Statistical Mechanics and its Applications* 418 (2015), pp. 94–104.

[30] D. P. Feldman and J. P. Crutchfield. 'Measures of statistical complexity: Why?' In: *Physics Letters A* 238.4-5 (1998), pp. 244–252.

[31] P. Flandrin. *Time-frequency/time-scale analysis*. Vol. 10. Academic press, 1998.

[32] G. Goerg and C. Shalizi. 'Mixed LICORS: A nonparametric algorithm for predictive state reconstruction'. In: *Artificial Intelligence and Statistics*. 2013, pp. 289–297.

[33] R. M. Gray. *Entropy and information theory*. Springer Science & Business Media, 2011.

[34] H. Janicke, A. Wiebel, G. Scheuermann and W. Kollmann. 'Multifield visualization using local statistical complexity'. In: *IEEE Transactions on Visualization and Computer Graphics* 13.6 (2007), pp. 1384–1391.

[35] L. P. Kadanoff. *Statistical physics, statics, dynamics & renormalization*. World Scie,tific, 2000.

[36] M. Martin, A. Plastino and O. Rosso. 'Generalized statistical complexity measures: Geometrical and analytical properties'. In: *Physica A: Statistical Mechanics and its Applications* 369.2 (2006), pp. 439–462.

[37] J. Nadal and P. E. Grassberger. *From Statistical Physics to Statistical Inference and Back*. New York Heidelberg Berlin: Springer, 1994. URL: http://www.springer.com/physics/complexity/book/978-0-7923-2775-2.

[38] C. R. Shalizi and K. L. Shalizi. 'Blind construction of optimal nonlinear recursive predictors for discrete sequences'. In: *Proceedings of the 20th conference on Uncertainty in artificial intelligence*. AUAI Press. 2004, pp. 504–511.

[39] C. R. Shalizi, K. L. Shalizi and R. Haslinger. 'Quantifying Self-Organization with Optimal Predictors'.
      In: *Phys. Rev. Lett.* 93 (11 Sept. 2004), p. 118701. DOI: 10.1103/PhysRevLett.93.118701. URL:
      https://link.aps.org/doi/10.1103/PhysRevLett.93.118701.

[40] J. L. Starck, F. Murthag and J. Fadili. *Sparse Image and Signal Processing: Wavelets, Curvelets,
      Morphological Diversity*. ISBN:9780521119139. Cambridge University Press, 2010.

[41] F. Takens. 'Detecting Strange Attractors in Turbulence'. In: *Non Linear Optimization* 898 (1981),
      pp. 366–381. URL: http://www.springerlink.com/content/b254x77553874745/.

[42] V. Venugopal and et al. 'Revisting multifractality of high resolution temporal rainfall using a wavelet-
      based formalism'. In: *Water Resources Research* 42 (2006).